BIOMARKER-SURROGATE (BIOSURROGATE) EVALUATION SCHEMA (BSES)

BIOMARKER-SURROGATE DOMAINS

Study Design Domain
0 Biological plausibility & lower quality clinical studies e.g. cross-sectional observational studies
1 Rank 0 and at least 2 good quality prospective observational cohort studies measuring S and T#
2 Rank 1 and at least 2 high quality adequately powered RCTs measuring S and T
3 Rank 1 and all, and at least 5 high quality adequately powered, RCTs measuring S and T

Target Outcome Domain
0 Target is reversible disease-centred biomarker of harm
1 Target is irreversible disease-centred biomarker of harm
2 Target is patient-centred endpoint of reversible organ morbidity or clinical burden of disease or clinical harm
3 Target is patient-centred endpoint of irreversible organ morbidity or clinical burden of disease or severe irreversible clinical harm or death

Statistical Evaluation of BioSurrogate – Target (B-T) Domain
0 Poor: Does not meet the criteria for Rank 1
1 Fair: RCT $R^2_{val} \geq 0.2$ AND STEP* $\geq 0.1$ AND $R^2_{ind} \geq 0.2$ OR cohort data $R^2_{val} \geq 0.4$
2 Good: RCT $R^2_{val} \geq 0.4$ AND STEP $\geq 0.2$ AND $R^2_{ind} \geq 0.4$
3 Excellent: RCT $R^2_{val} \geq 0.6$ AND STEP $\geq 0.3$ AND $R^2_{ind} \geq 0.6$ (without data subdivision)**

Generalisability of BioSurrogate-Target Domain:
Clinical evidence across different risk populations & pharmaco logic evidence across different drug-class mechanisms
0 No clinical or pharmacologic evidence
1 Clinical OR pharmacologic evidence
2 Clinical AND pharmacologic evidence
3 Consistent Clinical RCT AND pharmacologic RCT evidence

# Where S is the surrogate / biomarker/ biosurrogate and T is the target / true outcome
* STEP is defined as that proportion of the total range of the surrogate that is equal or larger than the STE
** Some analyses with few trials subdivide into centres to increase the number of data points

LEVEL OF EVIDENCE OF SURROGATE ENDPOINT MULTIDIMENSIONAL VALIDITY

A high rank on any one or more domain should not be allowed to prevail over a low rank on one or more domain when determining the overall level of evidence because at least good evidence of surrogacy across all domains is needed for surrogate validity. An A, B+, B, B- level surrogate endpoint ranks at least 2 on all domains.

Steps to determine the level of evidence:
1. The one and the same ‘evidence-base’ is applied across all four domains when determining the level of evidence.
2. Sum the highest rankings achieved across the four domains.
3. If any one domain is less than Rank 2, the level of evidence drops by one alphabetic category irrespective of the initial level. For example, B becomes a C, B- becomes a C- , C- becomes a D- and so forth.

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>level A</td>
</tr>
<tr>
<td>11-9</td>
<td>level B+, B, B-</td>
</tr>
<tr>
<td>8-6</td>
<td>level C+, C, C-, D+, D, D-</td>
</tr>
<tr>
<td>5-3</td>
<td>level D+, D, D-, E+, E, E-</td>
</tr>
<tr>
<td>2-0</td>
<td>level E+, E, E- F+, F, F-</td>
</tr>
</tbody>
</table>

Figure 1. Biomarker-Surrogate (BioSurrogate) Evaluation Schema (BSES2011)